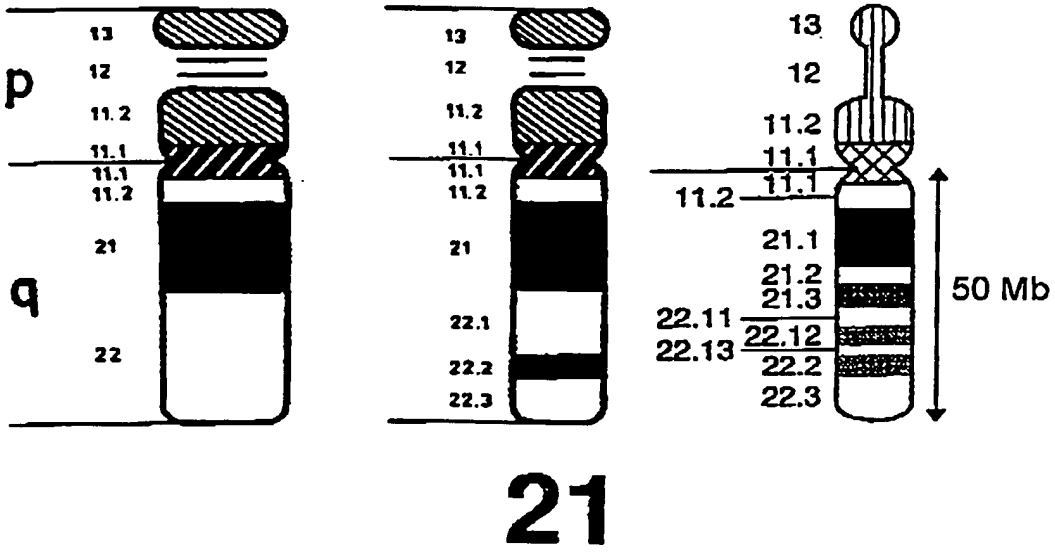



Figure 1

09850203 - 055501

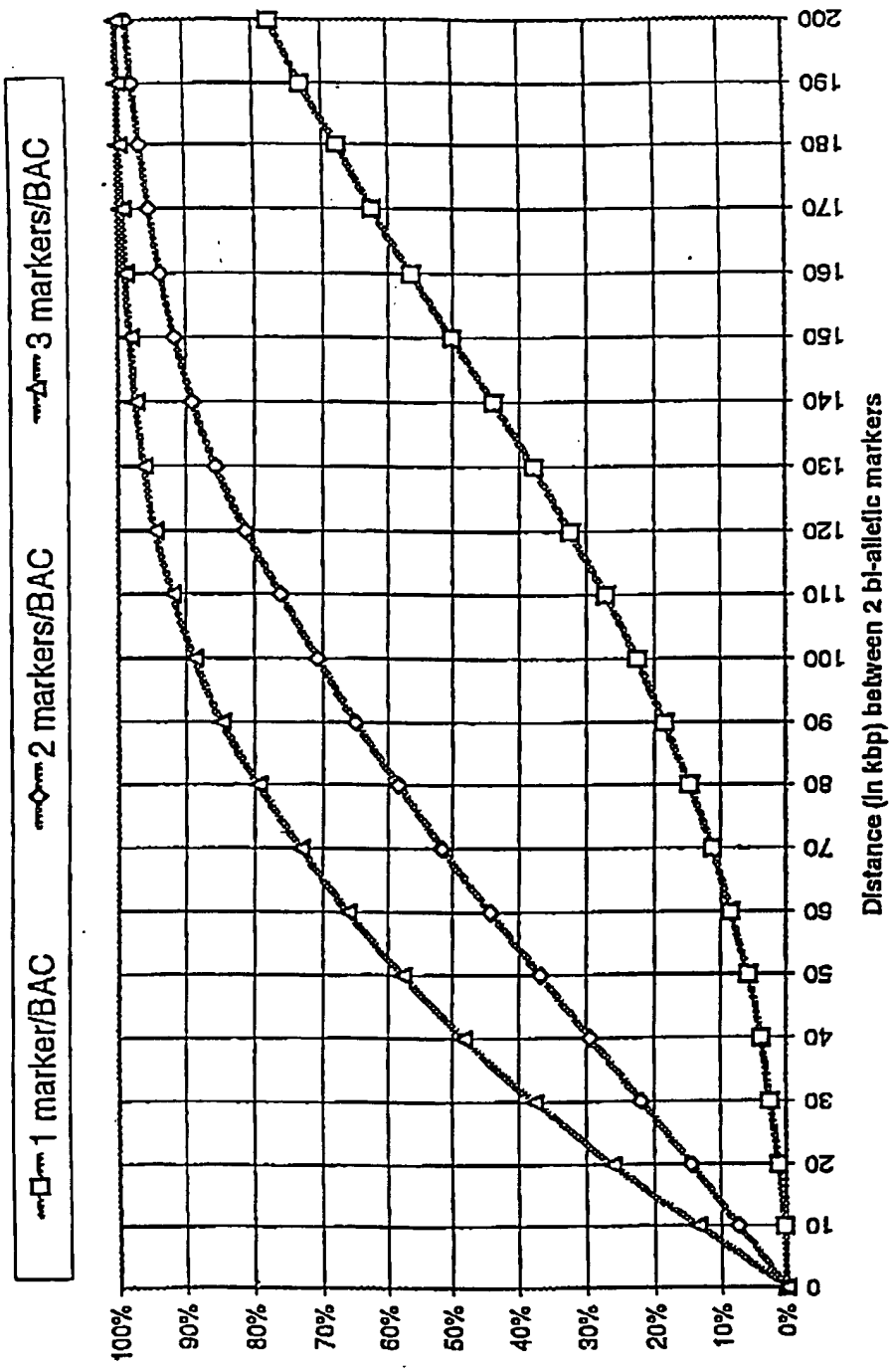



Figure 2A

FOSTER - 68258860

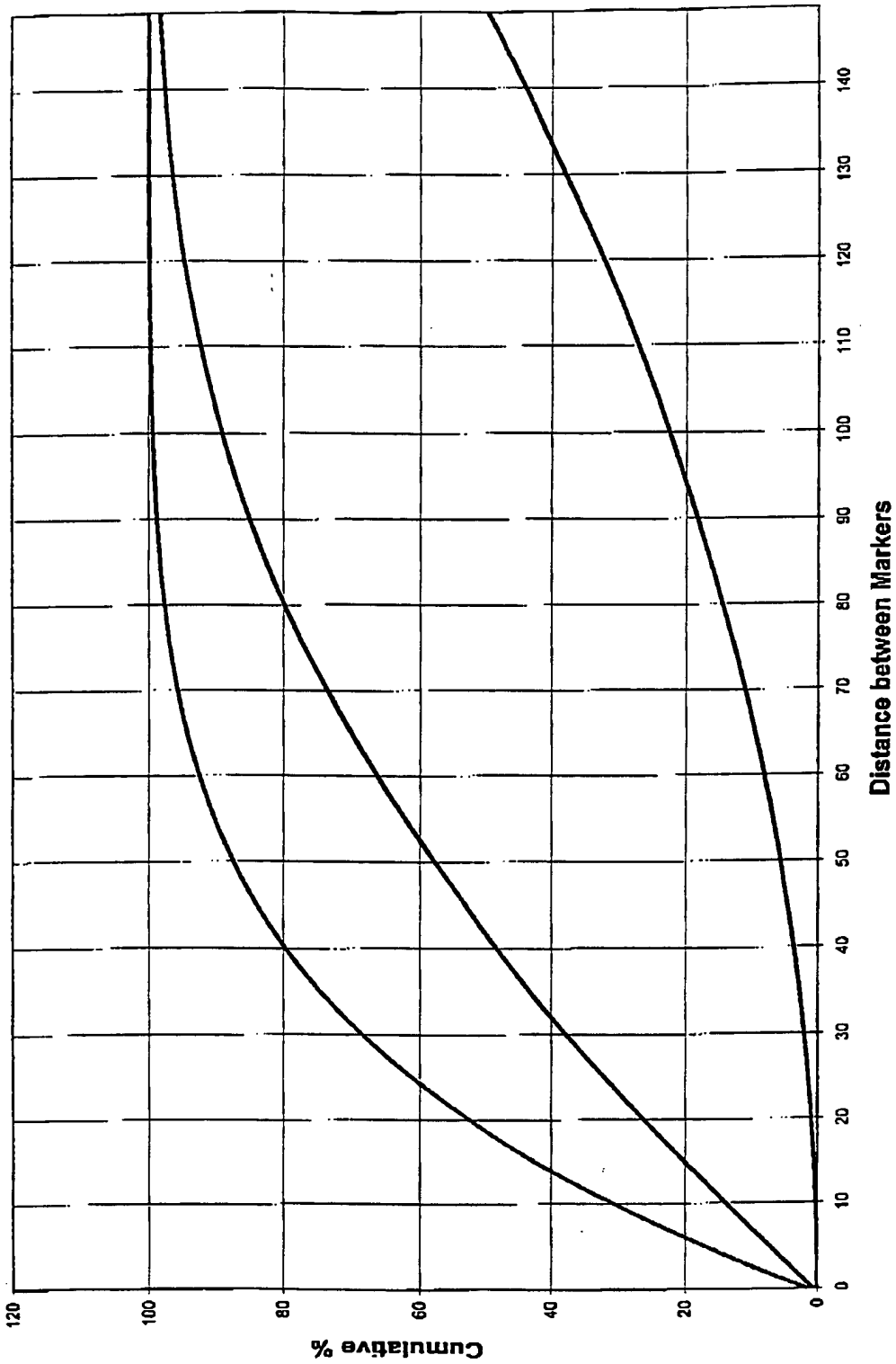
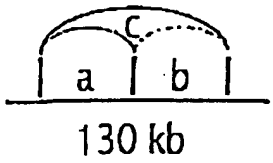



Figure 2B

09852830 - 051501

# LD in a random French caucasian population

- 54 sized « random » BACs covering 8100 kb
- 213 SNP ; 2 to 6 / BAC, mean allele frequency = 0.3
- Order and distance unknown
- For 1 BAC :




  * $\overline{m}$ intermarker distance : 130/3 = 43 kb
  * $\overline{m}$ LD strength estimate : m (a,b,c) = 0.51

- For 54 BACs :
  * $\overline{m}$ intermarker distance = 38 kb
  * $\overline{m}$ LD strength estimate = 0.63 ± 0.05 (324 pairs)
- For 19 unlinked SNPs : m LD strength estimate = 0.12 ± 0.007 (171 pairs)

Figure 2c

0852690.0501

p-VALUE DISTRIBUTION

| # aff | 150 | | | | | | |
|---|---|---|---|---|---|---|---|
| # non aff | 150 | | | | | | |
| | pAi non aff | 0 | 0,1 | 0,2 | 0,3 | 0,4 | 0,5 |
| Δ pAi | 0,05 | 8,77E-05 | 0,06407752 | 0,14252002 | 0,19106311 | 0,21543442 | 0,22009395 |
| Δ pAi | 0,1 | 1,91E-08 | 0,00060364 | 0,00467774 | 0,01023571 | 0,01382303 | 0,01382303 |
| Δ pAi | 0,15 | 3,06E-12 | 1,3319E-06 | 3,8827E-05 | 0,0001478 | 0,0002343 | 0,00020218 |
| Δ pAi | 0,2 | 3,22E-16 | 9,1413E-10 | 9,0305E-08 | 5,733E-07 | 9,6336E-07 | 5,733E-07 |
| Δ pAi | 0,25 | 2,08E-20 | 2,2614E-13 | 6,2679E-11 | 5,873E-10 | 8,7113E-10 | 2,5396E-10 |
| Δ pAi | 0,3 | 7,82E-25 | 2,152E-17 | 1,3261E-14 | 1,5189E-13 | 1,5189E-13 | 1,3261E-14 |
| Δ pAi | 0,35 | 1,62E-29 | 7,9823E-22 | 8,4152E-19 | 9,1669E-18 | 4,2713E-18 | 5,5844E-20 |
| Δ pAi | 0,4 | 1,73E-34 | 1,1282E-26 | 1,524E-23 | 1,1488E-22 | 1,524E-23 | 1,1282E-26 |

| # aff | 200 | | | | | | |
|---|---|---|---|---|---|---|---|
| # non aff | 200 | | | | | | |
| | pAi non aff | 0 | 0,1 | 0,2 | 0,3 | 0,4 | 0,5 |
| Δ pAi | 0,05 | 5,92E-06 | 0,03250945 | 0,09039173 | 0,13111935 | 0,15260313 | 0,15678006 |
| Δ pAi | 0,1 | 8,65E-11 | 7,4765E-05 | 0,00109084 | 0,00302686 | 0,00447365 | 0,00447365 |
| Δ pAi | 0,15 | 8,02E-16 | 2,3653E-08 | 2,0257E-06 | 1,1771E-05 | 2,1573E-05 | 1,7772E-05 |
| Δ pAi | 0,2 | 4,18E-21 | 1,5375E-12 | 6,7374E-10 | 7,764E-09 | 1,5417E-08 | 7,764E-09 |
| Δ pAi | 0,25 | 1,13E-26 | 2,525E-17 | 4,4025E-14 | 8,5532E-13 | 1,4423E-12 | 2,8149E-13 |
| Δ pAi | 0,3 | 1,47E-32 | 1,1488E-22 | 5,8424E-19 | 1,4886E-17 | 1,4886E-17 | 5,8424E-19 |
| Δ pAi | 0,35 | 8,62E-39 | 1,4784E-28 | 1,5457E-24 | 3,6958E-23 | 1,3394E-23 | 4,197E-26 |
| Δ pAi | 0,4 | 2,09E-45 | 5,2308E-35 | 7,6438E-31 | 1,1224E-29 | 7,6438E-31 | 5,2308E-35 |

| | |
|---|---|
| # aff | affected individuals |
| # non aff | non affected individuals |
| pAi non aff | allele frequency in non affected individuals |
| Δ pAi | % Difference in allele frequency between affected and non-affected individuals |

**Figure 3 (I)**

Title: METHODS, SOFTWARE AND APP... FOR IDENTIFYING GENOMIC REGIONS HAR... A GENE ASSOCIATED WITH A DETECTABL... TRAIT
Inventors: Nicholas SCHORK, et al.
Assignee: Genset Corporation
Our Ref.: 55.US4.DIV
5/31

p-VALUE DISTRIBUTION

| # aff | 500 | | | | | | |
|---|---|---|---|---|---|---|---|
| # non aff | 500 | | | | | | |
| | pAi non aff | 0 | 0,1 | 0,2 | 0,3 | 0,4 | 0,5 |
| Δ pAi | 0,05 | 8E-13 | 0,00072323 | 0,00741965 | 0,0169842 | 0,02371865 | 0,02516449 |
| Δ pAi | 0,1 | 1,07E-24 | 3,7948E-10 | 2,4176E-07 | 2,7579E-06 | 6,9679E-06 | 6,9679E-06 |
| Δ pAi | 0,15 | 3,81E-37 | 1,0719E-18 | 5,8344E-14 | 4,2622E-12 | 1,8601E-11 | 1,1611E-11 |
| Δ pAi | 0,2 | 2,96E-50 | 5,0895E-29 | 1,6881E-22 | 6,9321E-20 | 3,7441E-19 | 6,9321E-20 |
| Δ pAi | 0,25 | 4,27E-64 | 7,2043E-41 | 7,7528E-33 | 1,194E-29 | 4,3462E-29 | 7,6438E-31 |
| Δ pAi | 0,3 | 9,7E-79 | 3,9328E-54 | 6,3017E-45 | 1,9429E-41 | 1,9429E-41 | 6,3017E-45 |
| Δ pAi | 0,35 | 2,91E-94 | 8,8513E-69 | 8,7879E-59 | 2,3478E-55 | 1,8839E-56 | 1,1206E-62 |
| Δ pAi | 0,4 | 9,5E-111 | 7,7199E-85 | 1,8063E-74 | 1,4484E-71 | 1,8063E-74 | 7,7199E-85 |

| # aff | 150 | | | | | | |
|---|---|---|---|---|---|---|---|
| # non aff | 850 | | | | | | |
| | pAi non aff | 0 | 0,1 | 0,2 | 0,3 | 0,4 | 0,5 |
| Δ pAi | 0,05 | 2,16E-20 | 0,00994614 | 0,04896055 | 0,08358651 | 0,10417953 | 0,11025423 |
| Δ pAi | 0,1 | 2,01E-39 | 5,571E-07 | 0,00010149 | 0,00058665 | 0,00119145 | 0,00139743 |
| Δ pAi | 0,15 | 1,11E-58 | 2,7555E-13 | 8,462E-09 | 2,9851E-07 | 1,2395E-06 | 1,6229E-06 |
| Δ pAi | 0,2 | 3,27E-78 | 2,1683E-21 | 3,2211E-14 | 1,1049E-11 | 1,111E-10 | 1,5638E-10 |
| Δ pAi | 0,25 | 4,96E-98 | 4,4952E-31 | 6,5226E-21 | 3,1015E-17 | 2,5169E-16 | 1,1763E-15 |
| Δ pAi | 0,3 | 3,7E-118 | 3,6987E-42 | 8,129E-29 | 6,9335E-24 | 5,4331E-22 | 6,5657E-22 |
| Δ pAi | 0,35 | 1,4E-138 | 1,6797E-54 | 7,1058E-38 | 1,2938E-31 | 2,9415E-29 | 2,5869E-29 |
| Δ pAi | 0,4 | 2,4E-159 | 5,4915E-68 | 4,8846E-48 | 2,1003E-40 | 1,3332E-37 | 6,8178E-38 |

| | |
|---|---|
| # aff | affected individuals |
| # non aff | non affected individuals |
| pAi non aff | allele frequency in non affected individuals |
| Δ pAi | % Difference in allele frequency between affected and non-affected individuals |

**Figure 3 (II)**

p-VALUE DISTRIBUTION

| # aff | 200 | | | | | | |
|---|---|---|---|---|---|---|---|
| # non aff | 500 | | | | | | |
| | pAi non aff | 0 | 0,1 | 0,2 | 0,3 | 0,4 | 0,5 |
| Δ pAi | 0,05 | 1,06E-12 | 0,00789803 | 0,03942584 | 0,06867566 | 0,08621572 | 0,09083704 |
| Δ pAi | 0,1 | 3,45E-24 | 4,4217E-07 | 5,6883E-05 | 0,00031976 | 0,0006363 | 0,00070881 |
| Δ pAi | 0,15 | 5,9E-36 | 4,3025E-13 | 3,3635E-09 | 9,2134E-08 | 3,319E-07 | 3,5871E-07 |
| Δ pAi | 0,2 | 4,73E-48 | 1,5566E-20 | 1,0346E-14 | 1,7218E-12 | 1,1512E-11 | 1,0047E-11 |
| Δ pAi | 0,25 | 1,67E-60 | 3,5436E-29 | 2,0473E-21 | 2,2178E-18 | 1,1498E-17 | 1,3524E-17 |
| Δ pAi | 0,3 | 2,46E-73 | 7,2498E-39 | 3,0748E-29 | 2,0601E-25 | 3,4525E-24 | 7,4807E-25 |
| Δ pAi | 0,35 | 1,44E-86 | 1,6945E-49 | 3,9559E-38 | 1,4118E-33 | 2,662E-32 | 1,4118E-33 |
| Δ pAi | 0,4 | 3,2E-100 | 5,3051E-61 | 4,7325E-48 | 7,1282E-43 | 1,0691E-41 | 7,2652E-44 |

| # aff | 500 | | | | | | |
|---|---|---|---|---|---|---|---|
| # non aff | 1000 | | | | | | |
| | pAi non aff | 0 | 0,1 | 0,2 | 0,3 | 0,4 | 0,5 |
| Δ pAi | 0,05 | 6,48E-24 | 5,7827E-05 | 0,00172627 | 0,00551541 | 0,00882876 | 0,00978249 |
| Δ pAi | 0,1 | 6,53E-47 | 3,065E-14 | 1,0301E-09 | 4,3205E-08 | 1,8833E-07 | 2,2731E-07 |
| Δ pAi | 0,15 | 1,2E-70 | 2,0716E-27 | 3,7441E-19 | 4,6626E-16 | 6,9719E-15 | 6,9719E-15 |
| Δ pAi | 0,2 | 3,33E-95 | 1,1636E-43 | 1,6614E-31 | 8,5632E-27 | 4,1421E-25 | 1,9885E-25 |
| Δ pAi | 0,25 | 1,2E-120 | 1,7683E-62 | 1,5329E-46 | 3,1722E-40 | 8,6765E-39 | 3,6071E-39 |
| Δ pAi | 0,3 | 5,3E-147 | 1,526E-83 | 4,2697E-64 | 2,5968E-56 | 3,9328E-54 | 2,5968E-56 |
| Δ pAi | 0,35 | 2,4E-174 | 1,184E-106 | 4,5658E-84 | 4,7426E-75 | 4,2624E-73 | 4,0958E-77 |
| Δ pAi | 0,4 | 9,4E-203 | 1,082E-131 | 2,137E-106 | 1,8014E-96 | 3,3252E-95 | 6,725E-102 |

| | |
|---|---|
| # aff | affected individuals |
| # non aff | non affected individuals |
| pAi non aff | allele frequency in non affected individuals |
| Δ pAi | % Difference in allele frequency between affected and non-affected individuals |

**Figure 3 (III)**

Title: METHODS, SOFTWARE AND APPARATI FOR IDENTIFYING GENOMIC REGIONS H[redacted]ING A GENE ASSOCIATED WITH A DETEC[redacted]RAIT
Inventors: Nicholas SCHORK
Assignee: Genset Corporation
Our Ref.: 55.US4.DIV
7/31

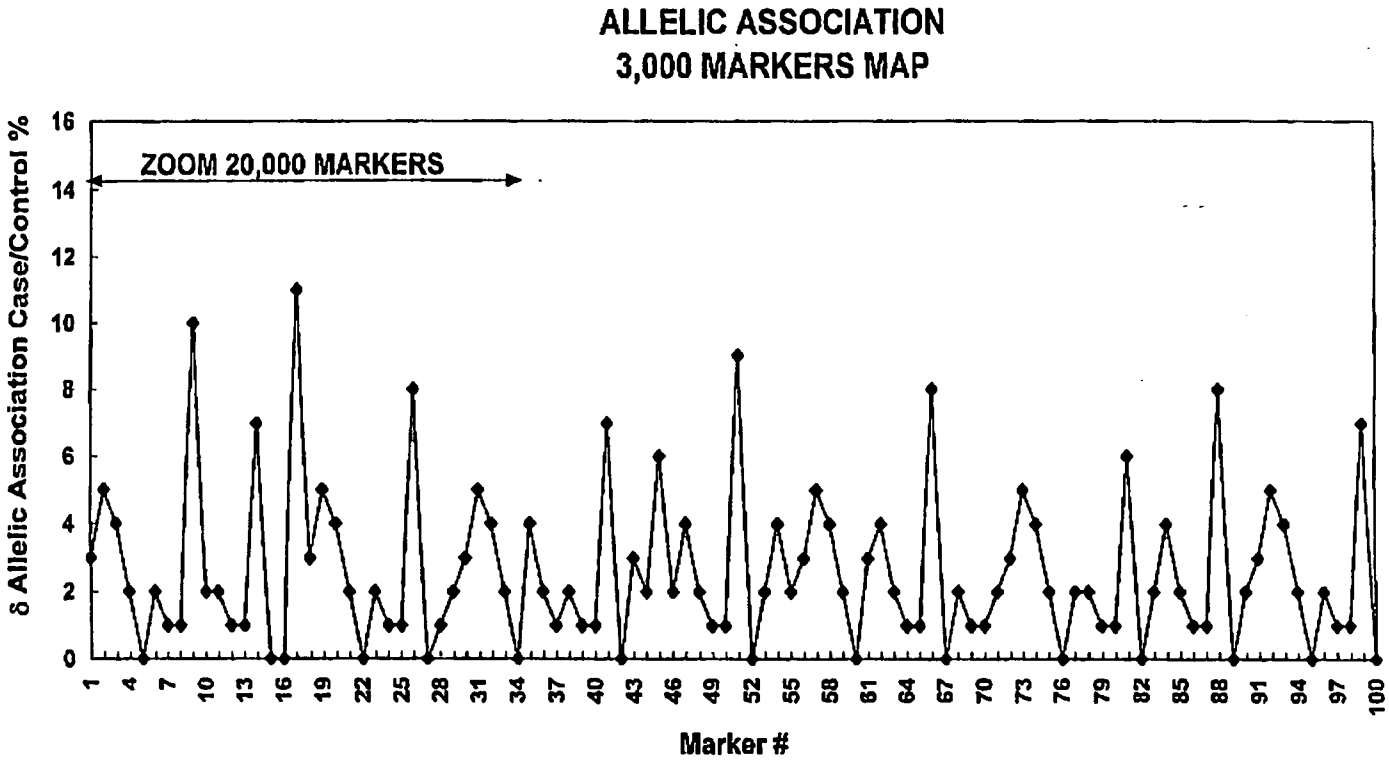



Figure 4

Title: METHODS, SOFTWARE AND ... FOR IDENTIFYING GENOMIC REGIONS ... A GENE ASSOCIATED WITH A DETECTABLE TRAIT
Inventors: Nicholas SCHORK, et al.
Assignee: Genset Corporation
Our Ref.: 55.US4.DIV
8/31

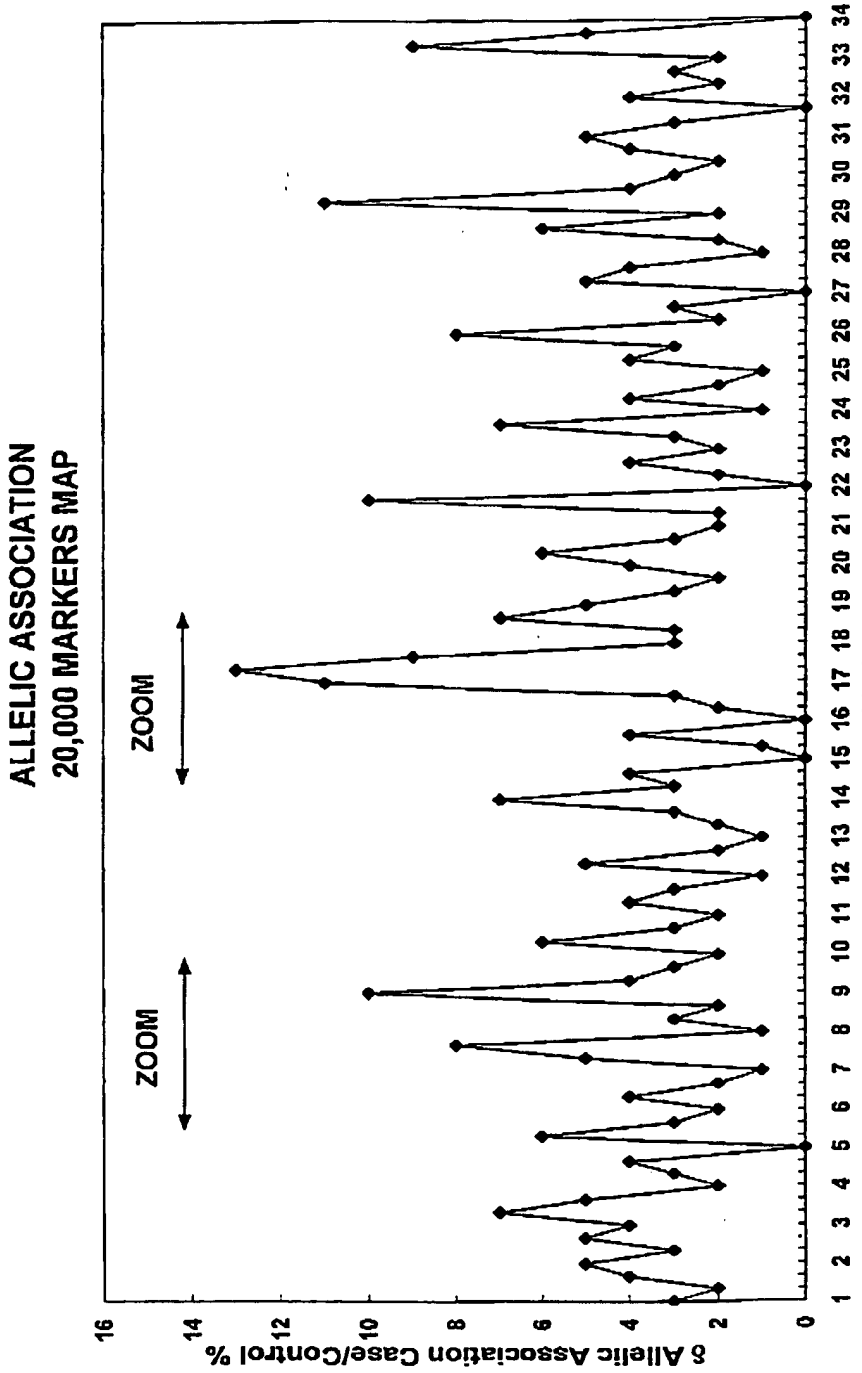



Figure 5

09532869 - 052501

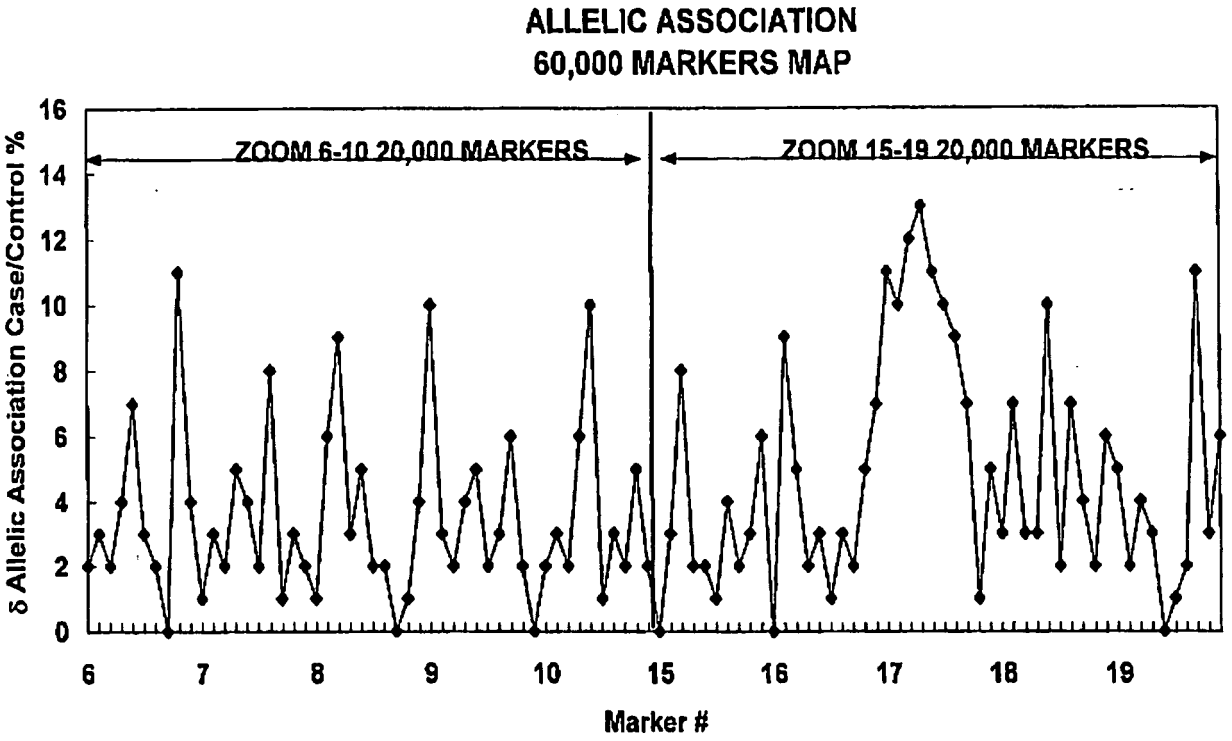



Figure 6

Title: METHODS, SOFTWARE AND APPARATI FOR IDENTIFYING GENOMIC REGIONS HARBORING A GENE ASSOCIATED WITH A DETECTABLE TRAIT
Inventors: Nicholas SCHORK, et al.
Assignee: Genset Corporation
Our Ref.: 55.US4.DIV
10/31

TOSTSO" 68285860

## APO E REGION HAPLOTYPE FREQUENCY ANALYSIS

| POPULATIONS | AD CASES (225) | AD CONTROLS (248) |
|---|---|---|

| markers | 99-366 | 99-344 | 99-359 | 99-355 | haplotype frequencies | | odds-ratio | P value |
|---|---|---|---|---|---|---|---|---|
| p value | 3,01E-01 | 1,11E-01 | 6,63E-01 | 1,38E-01 | cases | controls | | |
| haplotype 1 | C | G | | | 0,404 | 0,308 | 1,52 | **3,05E-03** *** |
| haplotype 2 | | G | A | | 0,203 | 0,165 | 1,29 | 1,24E-01 * |
| haplotype 3 | | | G | G | 0,375 | 0,306 | 1,36 | 2,83E-02 ** |
| haplotype 4 | C | | A | | 0,264 | 0,209 | 1,36 | 5,95E-02 ** |
| haplotype 5 | | G | | A | 0,116 | 0,071 | 1,70 | 1,64E-02 ** |
| haplotype 6 | C | | | A | 0,15 | 0,129 | 1,19 | 3,59E-01 * |
| haplotype 7 | T | | G | G | 0,225 | 0,122 | 2,09 | 4,76E-05 ***** |
| haplotype 8 | T | A | G | G | 0,228 | 0,108 | 2,44 | 2,05E-06 ****** |

Figure 7

Title: METHODS, SOFTWARE AND APPARATI FOR IDENTIFYING GENOMIC REGIONS HARBORING A GENE ASSOCIATED WITH A DETECTABLE TRAIT
Inventors: Nicholas SCHORK, et al.
Assignee: Genset Corporation
Our Ref.: 55.US4.DIV
11/31

# APO E REGION HAPLOTYPE SIMULATION
# POPULATION : 225 CASES vs 248 CONTROLS

| | 4 Markers | | | | haplotype frequencies | | odds-ratio | pvalue |
|---|---|---|---|---|---|---|---|---|
| | 99-344/439 | 99-366/274 | 99-359/308 | 99-355/219 | cases | controls | | |
| Haplotype 8 | A | T | G | G | 0,228 | 0,108 | 2,44 | 2,05E-06 ****** |

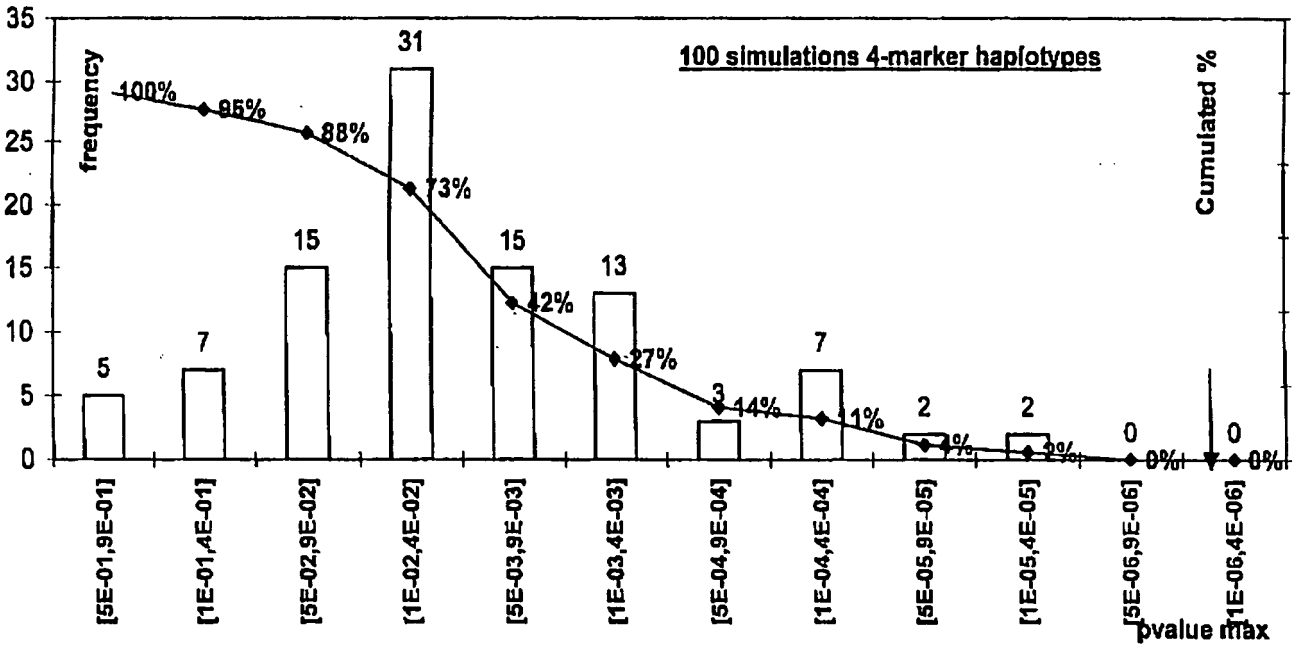



Figure 8

Title: METHODS, SOFTWARE AND A... IDENTIFYING GENOMIC REGIONS ... GENE ASSOCIATED WITH A DETEC... ...TRAIT
Inventors: Nicholas SCHORK, et al.
Assignee: Genset Corporation
Our Ref.: 55.US4.DIV
12/31

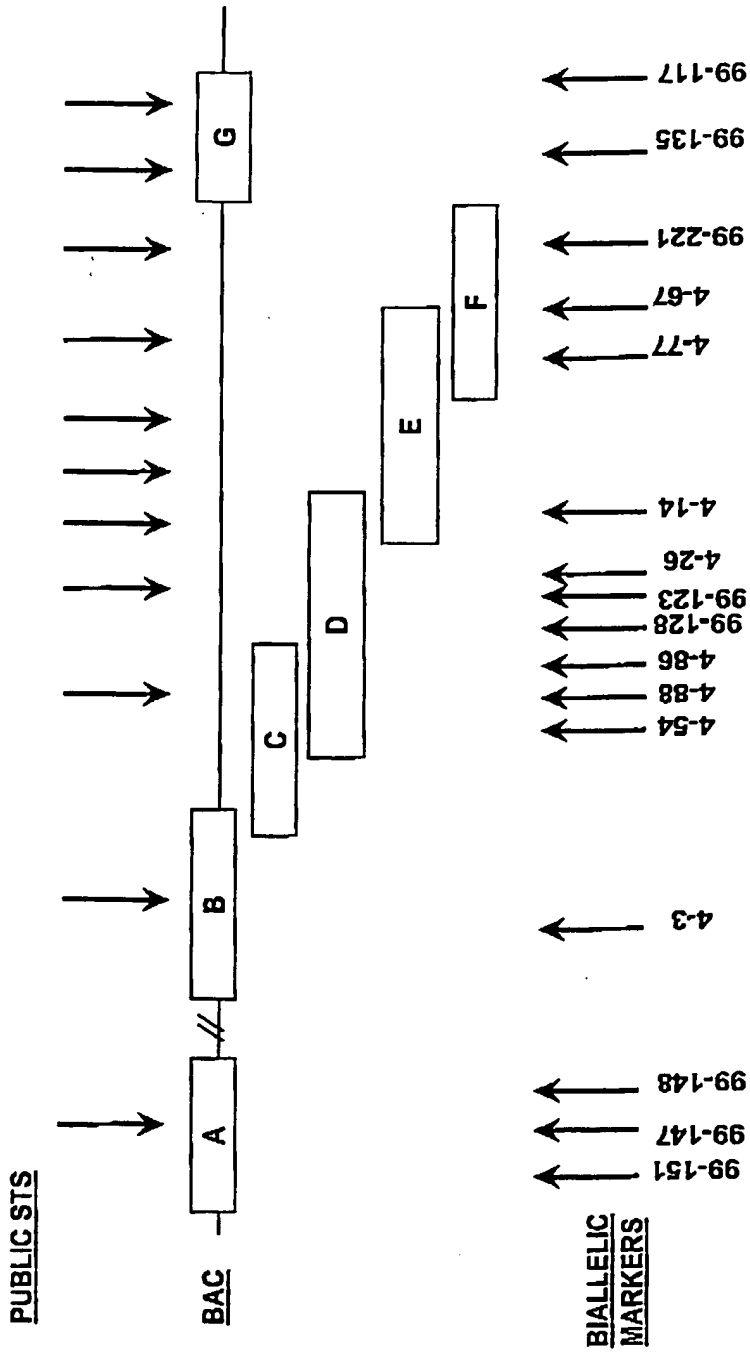



Figure 9

09852889 - 051501

# PROSTATE CANCER ASSOCIATION STUDIES (FIRST SCREENING)

| Population<br>Sample size | PROSTATE CANCER<br>CASES = 112 | NON AFFECTED<br>CONTROLS=76 |
|---|---|---|
| Population<br>Characteristics | 35 sporadic cases<br>+ 77 familial cases | > 65 years<br>PSA<4 |

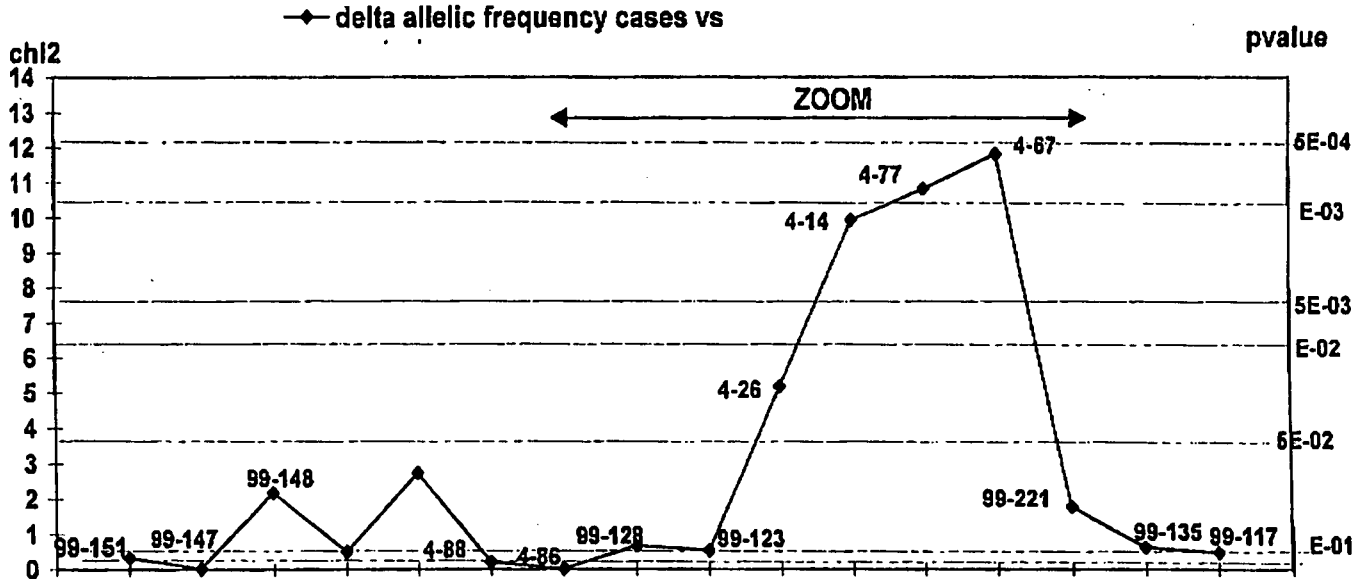



Figure 10

Title: METHODS, SOFTWARE AND APPARATI FOR IDENTIFYING GENOMIC REGIONS HARBORING A GENE ASSOCIATED WITH A DETEC[redacted]E TRAIT
Inventors: Nicholas SCHO[redacted]
Assignee: Genset Corpo[redacted]
Our Ref.: 55.US4.DIV
14/31

TOSTSO" 68285860

# PROSTATE CANCER ASSOCIATION STUDIES (ZOOM)

| | PROSTATE CANCER CASES (185) | NON-AFFECTED CONTROLS (104) |
|---|---|---|
| characteristics of populations | 47 sporadic cases + 138 familial cases | > 65 years PSA<4 |

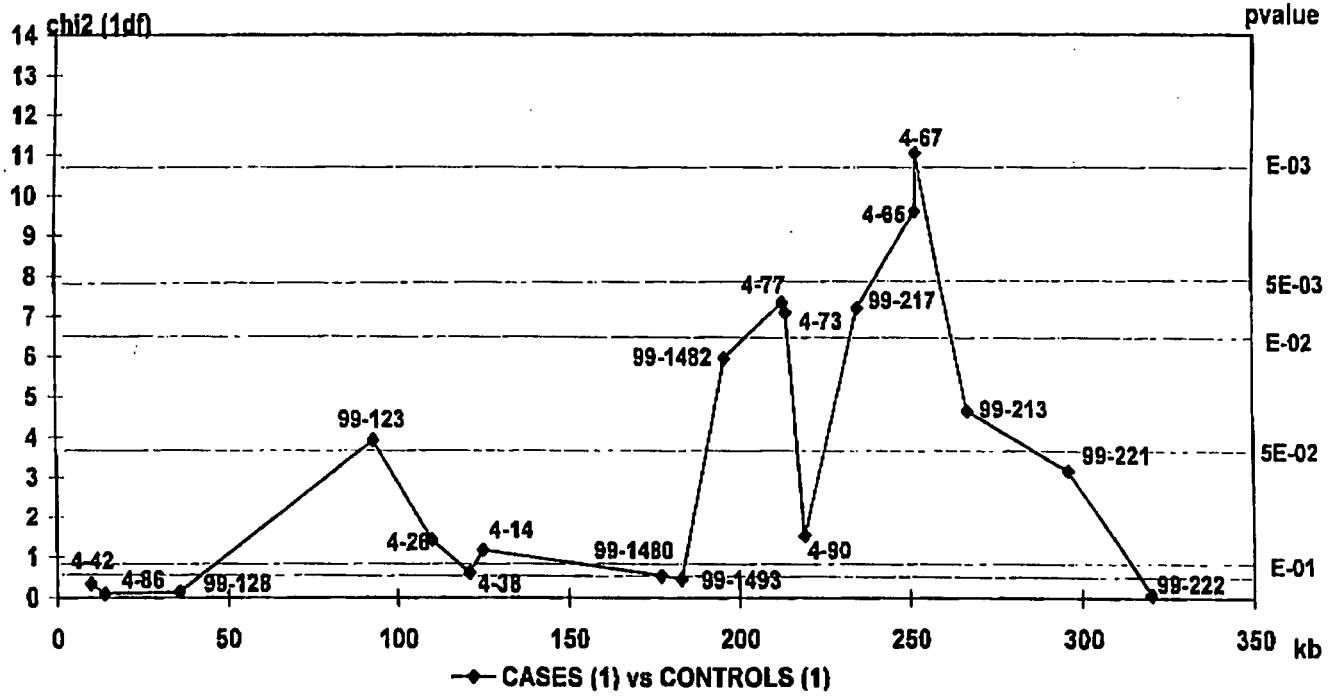



Figure 11

Title: METHODS, SOFTWARE AND APPARATI FOR IDENTIFYING GENOMIC REGIONS HARBORING A GENE ASSOCIATED WITH A DETECTABLE TRAIT
Inventors: Nicholas SCHORK
Assignee: Genset Corporation
Our Ref.: 55.US4.DIV
15/31

TOSTSO"6823S860



## PROSTATE CANCER HAPLOTYPE FREQUENCY ANALYSIS

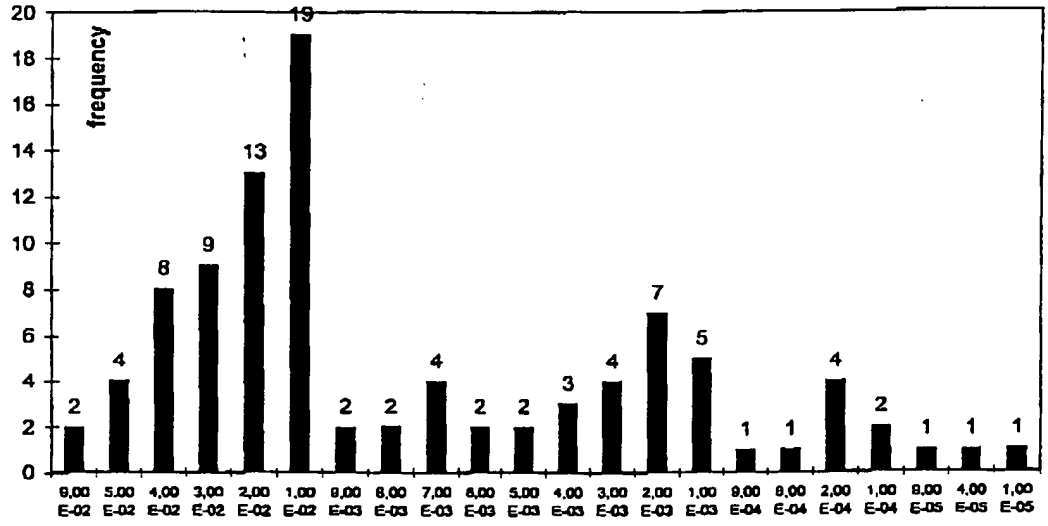
| | PROSTATE CANCER CASES (281) | NON-AFFECTED CONTROLS (130) |
|---|---|---|
| characteristics of populations | 143 sporadic cases + 138 familial cases | > 65 years PSA<4 |

| markers | 99-123 | 4-26 | 4-14 | 4-77 | 99-217 | 4-67 | 99-213 | 99-221 | 99-135 | haplotype frequencies | | relative risk | pvalue |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| bacs | H0287B09 | B0189E08 | | B0463F01 | | | | | B0725B12 | | | | |
| genes | <----------PG1----------> | | | | | | | | | | | | |
| p value | 2,00E-01 | 1,00E-01 | 1,00E-01 | 2,00E-02 | 2,00E-02 | 6,00E-04 | 9,00E-02 | 7,00E-01 | 2,00E-01 | cases | controls | | |
| haplotype 8 >304kb< | C | A | C | G | T | T | C | A | A | 0,075 | 0,018 | 4,42 | 9,00E-04 *** |
| haplotype 7 >286kb< | | A | C | G | T | T | C | A | A | 0,095 | 0,016 | 6,46 | 6,00E-05 **** |
| haplotype 6 <186kb> | | A | C | G | T | T | C | A | | 0,116 | 0,019 | 6,78 | 1,00E-05 ***** |
| haplotype 5 <171kb> | | | C | G | T | T | C | A | | 0,117 | 0,013 | 10,06 | 9,00E-07 ******* |
| haplotype 4 <83kb> | | | | G | T | T | C | A | | 0,117 | 0,025 | 5,17 | 2,00E-05 ***** |
| haplotype 3.1 <54kb> | | | | | T | T | C | A | | 0,117 | 0,027 | 4,78 | 2,00E-05 ***** |
| haplotype 3.2 <54kb> | | | | G | T | T | C | | | 0,222 | 0,109 | 2,33 | 4,00E-05 ***** |
| haplotype 2.2 <39kb> | | | | G | T | T | | | | 0,251 | 0,134 | 2,17 | 2,00E-04 **** |
| haplotype 2 <32kb> | | | | | T | T | C | | | 0,226 | 0,112 | 2,32 | 1,00E-04 *** |
| haplotype 1.1 <17 kb> | | | | | T | T | | | | 0,256 | 0,148 | 2,01 | 3,00E-04 **** |
| haplotype 1.2 <15 kb> | | | | | | T | C | | | 0,233 | 0,129 | 2,05 | 6,00E-04 *** |

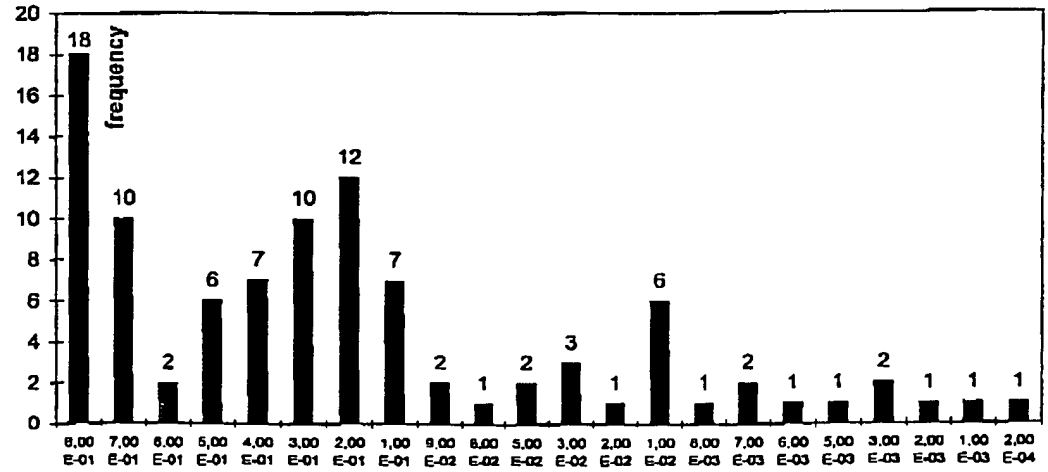
Figure 12

## PROSTATE CANCER HAPLOTYPE SIMULATIONS (100 ITERATIONS)

| | | | | | | | haplotype frequencies | | relative | pvalue |
|---|---|---|---|---|---|---|---|---|---|---|
| markers | 4-14 | 4-77 | 99-217 | 4-67 | 99-213 | 99-221 | cases | controls | risk | |
| haplotype | C | G | T | T | G | A | 0,117 | 0,013 | 10,06 | 9,00E-07 |




pvalue max of haplotypes for 100 simulations




pvalue of haplotype CGTTGA for 100 simulations

**Figure 13**

09852869 - 050450

# AVERAGE LD PATTERN GENOMIC HETEROGENEITY

| Recombination rate | Lower<br>A | Higher<br>B |
|---|---|---|
| Nb markers | 89 | 69 |
| All SNP | 0.61<br>(749) | 0.42<br>(1190) |
| Rare < 0.2<br>Rare vs rare | 0.75<br>(65) | 0.17<br>(158) |
| Frequent > 0.2<br>Frequent vs frequent | 0.51<br>(410) | 0.49<br>(544) |
| Rare vs frequent | 0.72<br>(274) | 0.41<br>(488) |

FIGURE 14

FR05T50.050502

## Exonic/nonexonic LD

| | Nb pairs | Average intermarker distance | Average LD |
|---|---|---|---|
| Exonic SNPs | 36 | 26 kb | 0.65±0.021 |
| Non exonic SNPs | 60 | 36 kb | 0.48±0.018 |
| Exonic/Non exonic | 96 | 32 kb | **0.60±0.015** |

**FIGURE 15**

D05E2003.055504

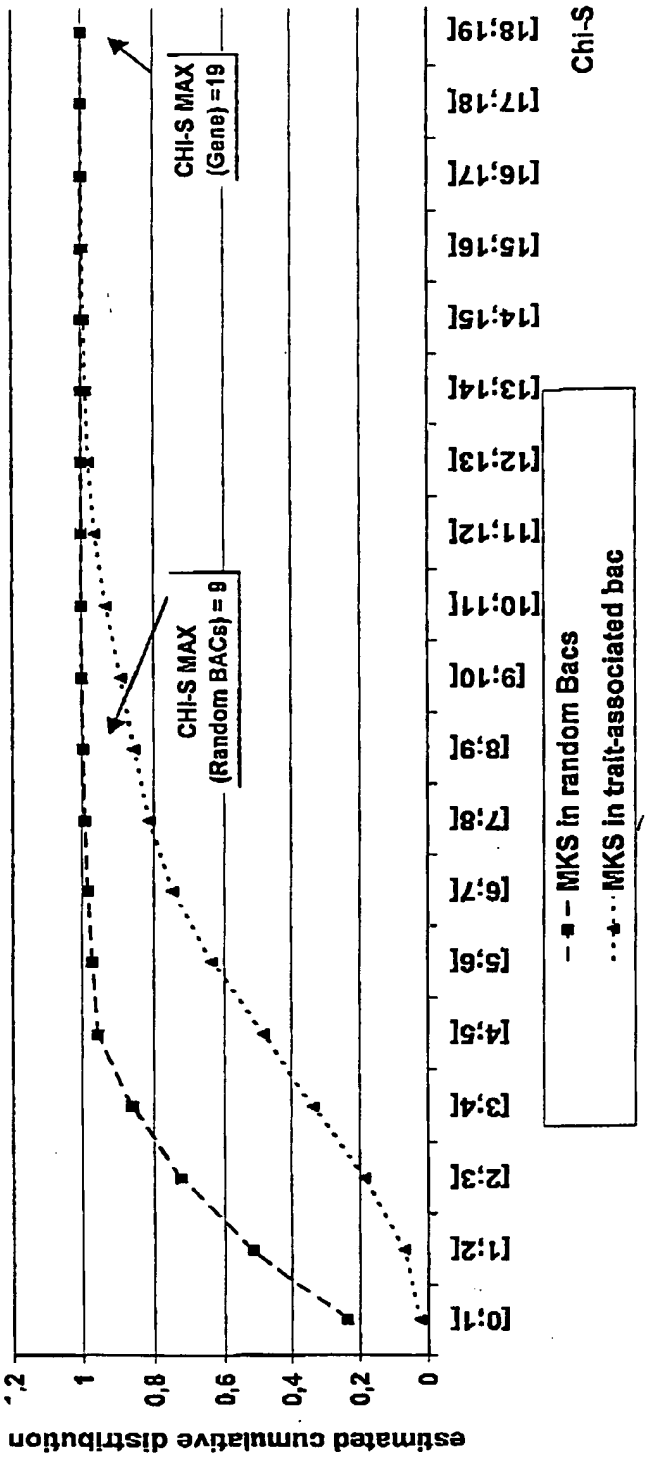



FIGURE 16 A

Patent - 55.US4.DIV

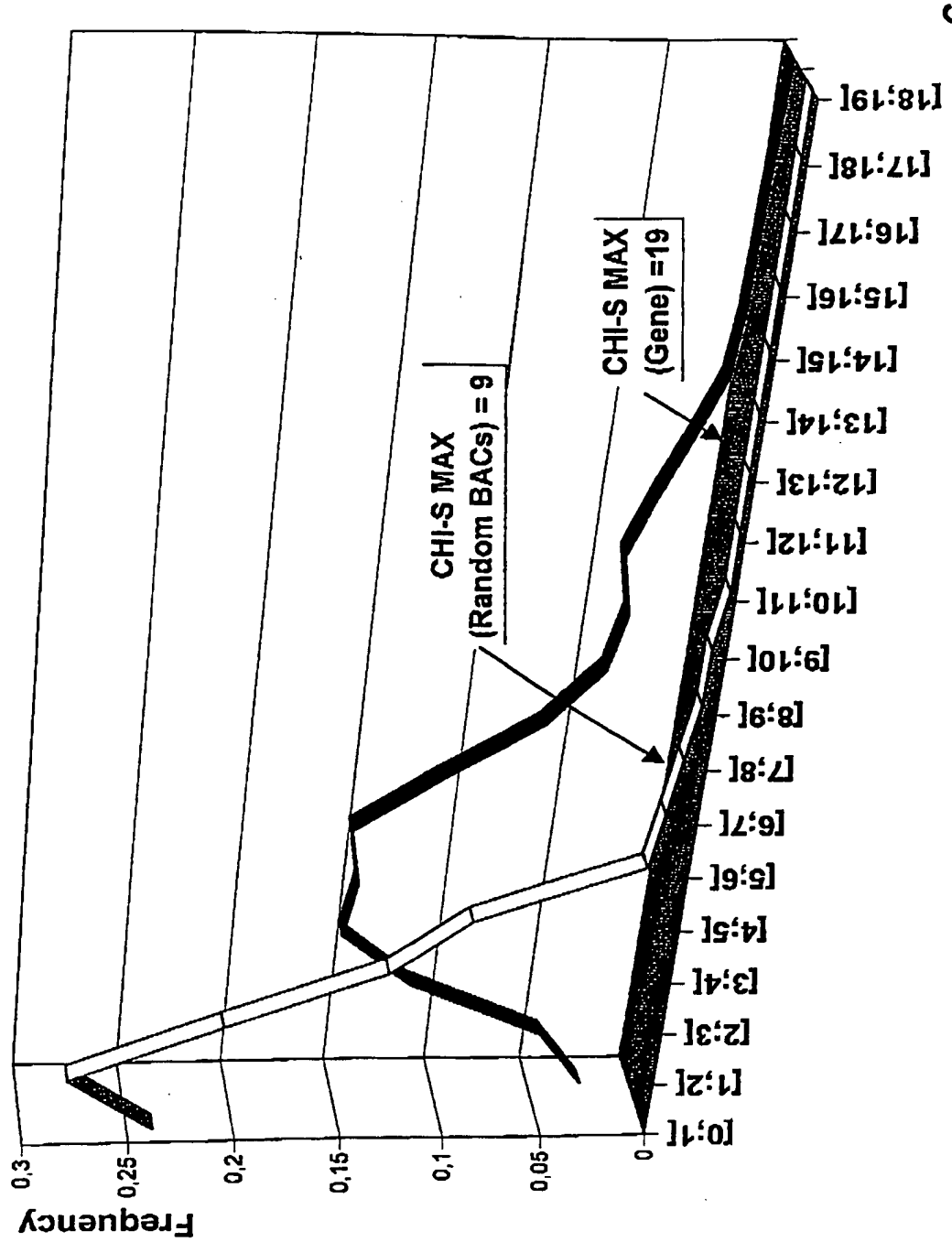



FIGURE 16B

FACTS0 - 0055400.0

TOSTSO"6828860

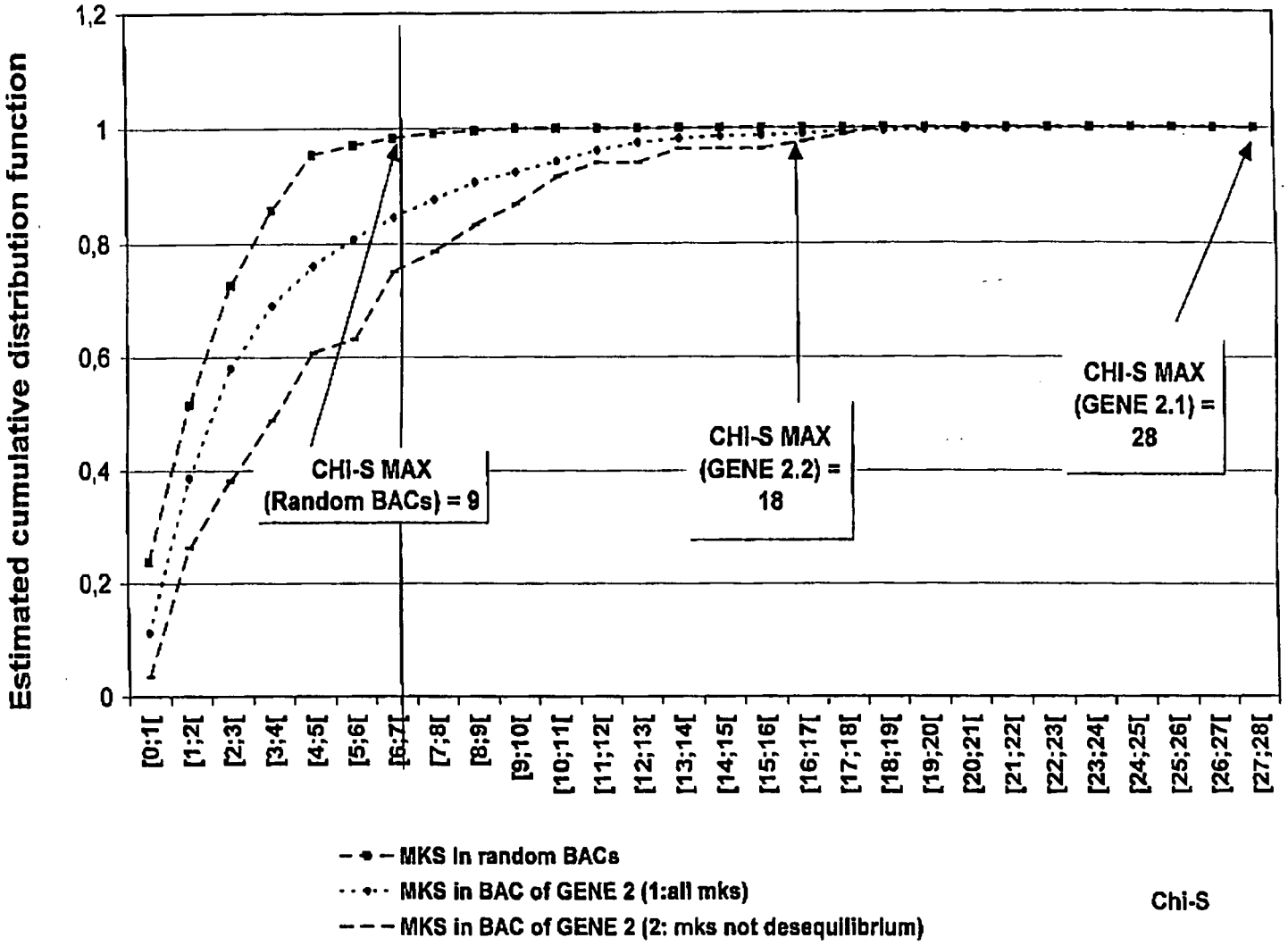



FIGURE 17A

Title: METHODS, SOFTWARE AND APPARATI FOR IDENTIFYING GENOMIC REGIONS HARBORING A GENE ASSOCIATED WITH A DETECTABLE TRAIT
Inventors: Nicholas SCHORK et al.
Assignee: Genset Corporation
Our Ref.: 55.US4.DIV
22/31

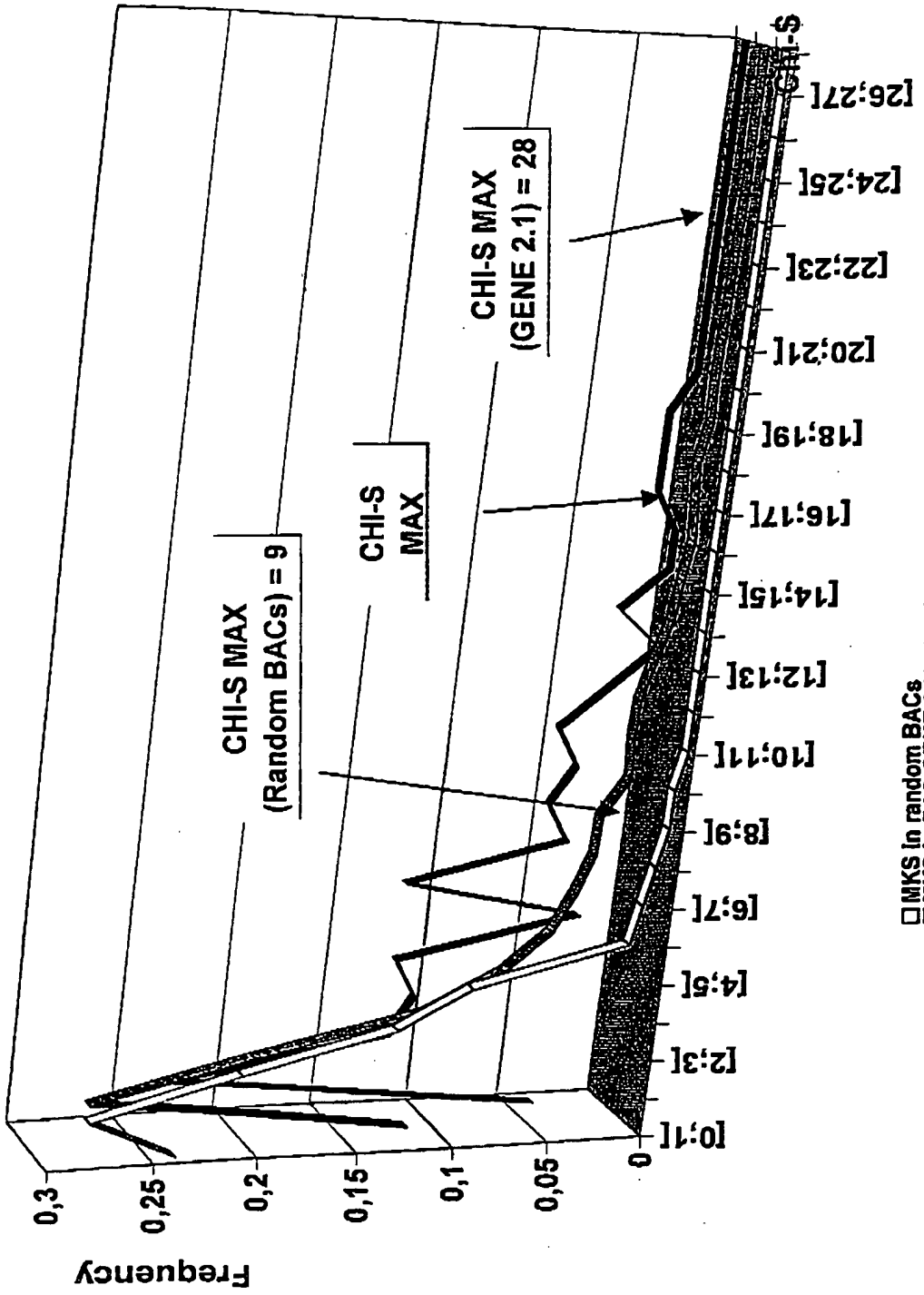

CHI-S MAX
(Random BACs) = 9
CHI-S
MAX
CHI-S MAX
(GENE 2.1) = 28
Frequency
0,3
0,25
0,2
0,15
0,1
0,05
0
[0;1[
[2;3[
[4;5[
[6;7[
[8;9[
[10;11[
[12;13[
[14;15[
[16;17[
[18;19[
[20;21[
[22;23[
[24;25[
[26;27[
MKS in random BACs
MKS in BAC GENE 2 (1:all mks)
MKS in BAC GENE 2 (2:not mks in desequilibrium)

FIGURE 17 B

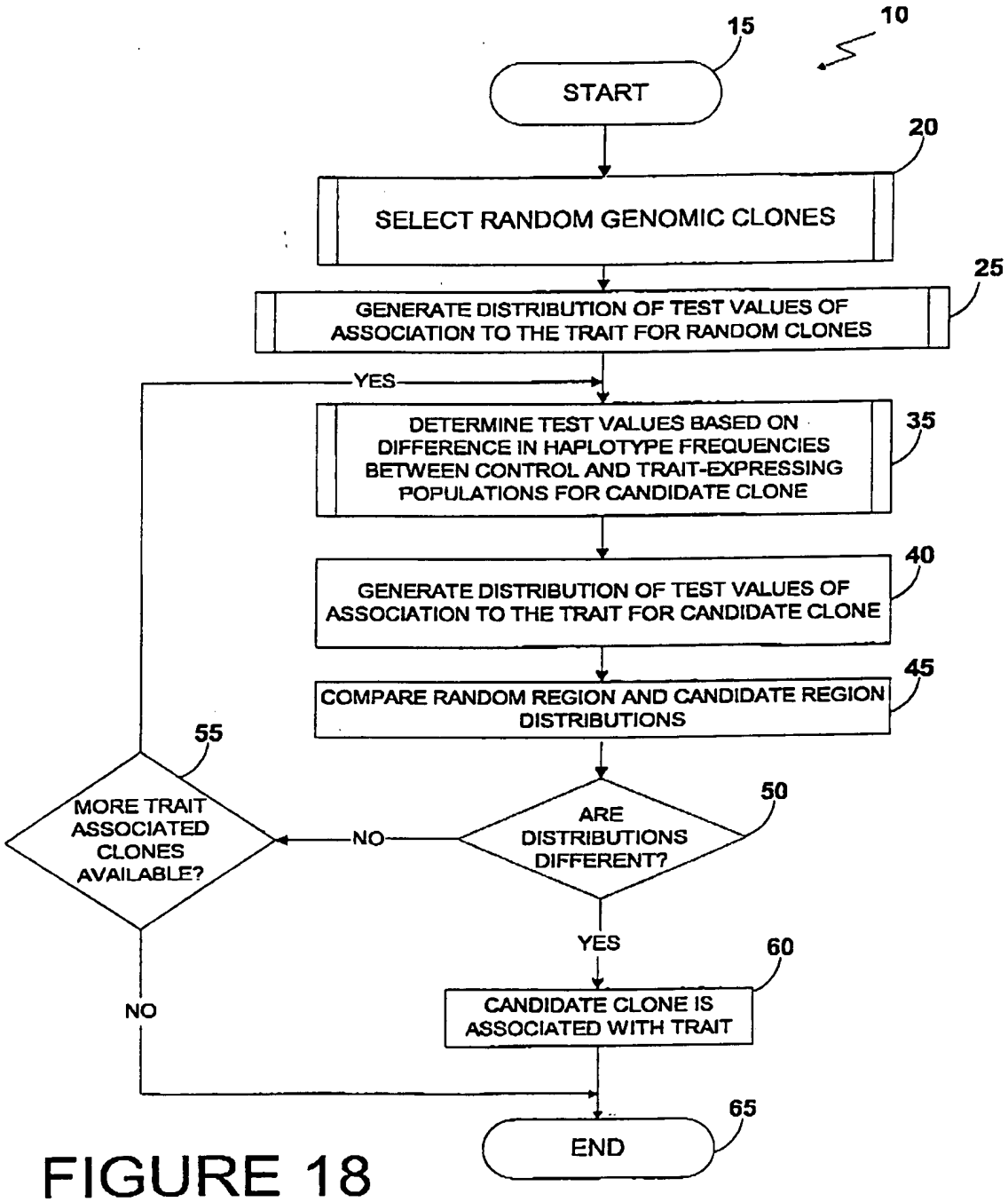



FIGURE 18

Docket No. 55.US4.DIV

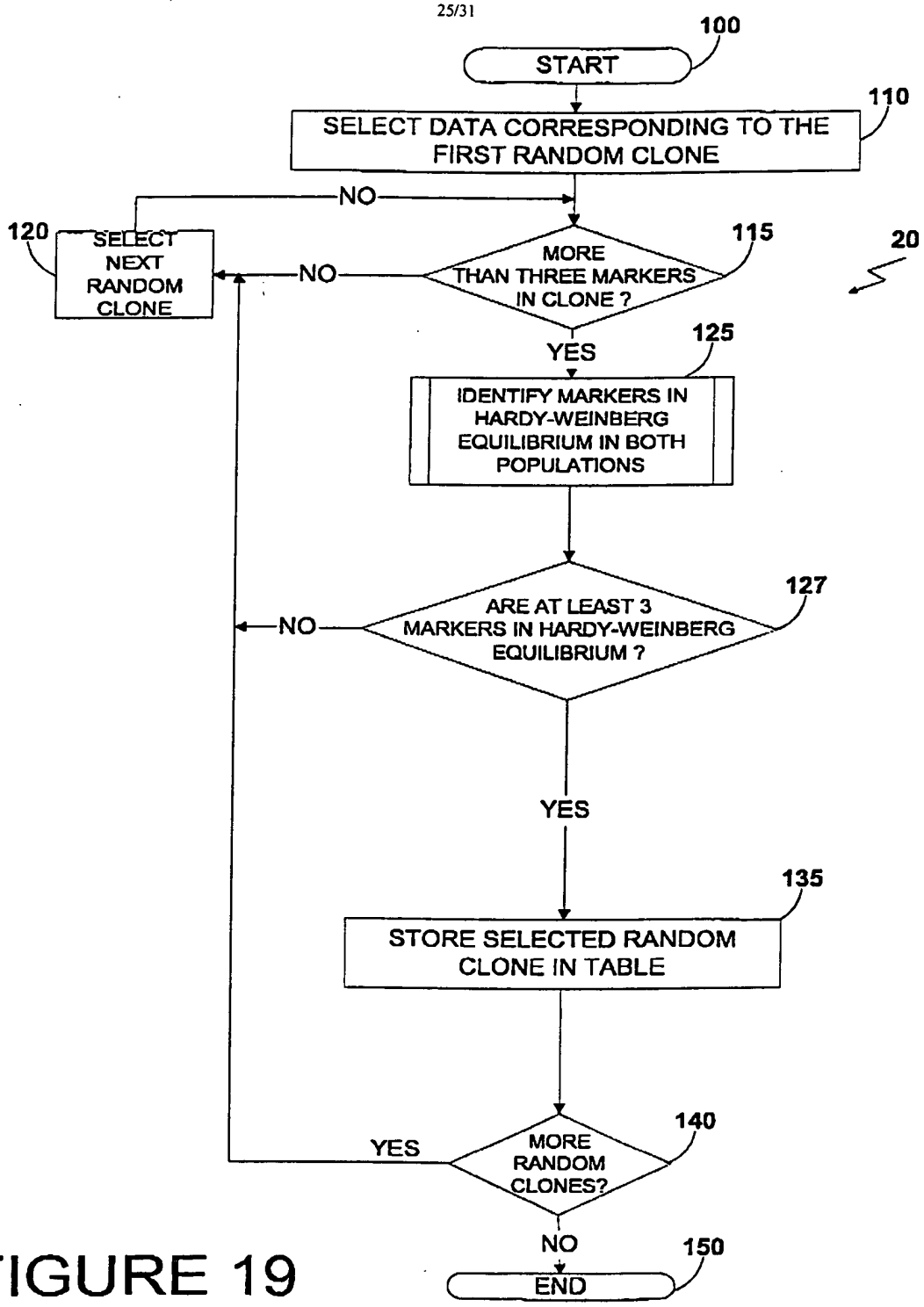



FIGURE 19

Docket No. 55.US4.DIV

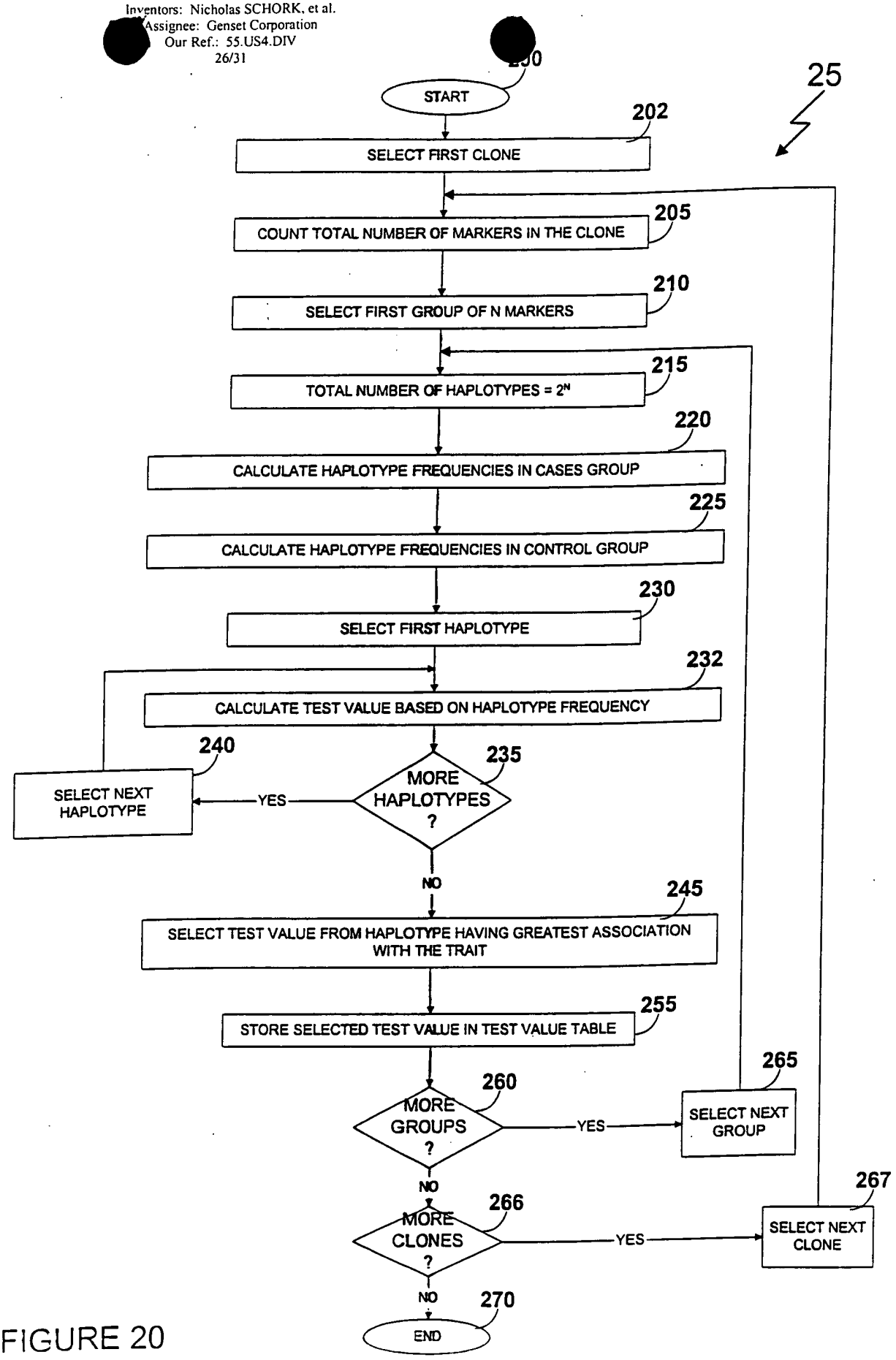



FIGURE 20

05050209 - 055450

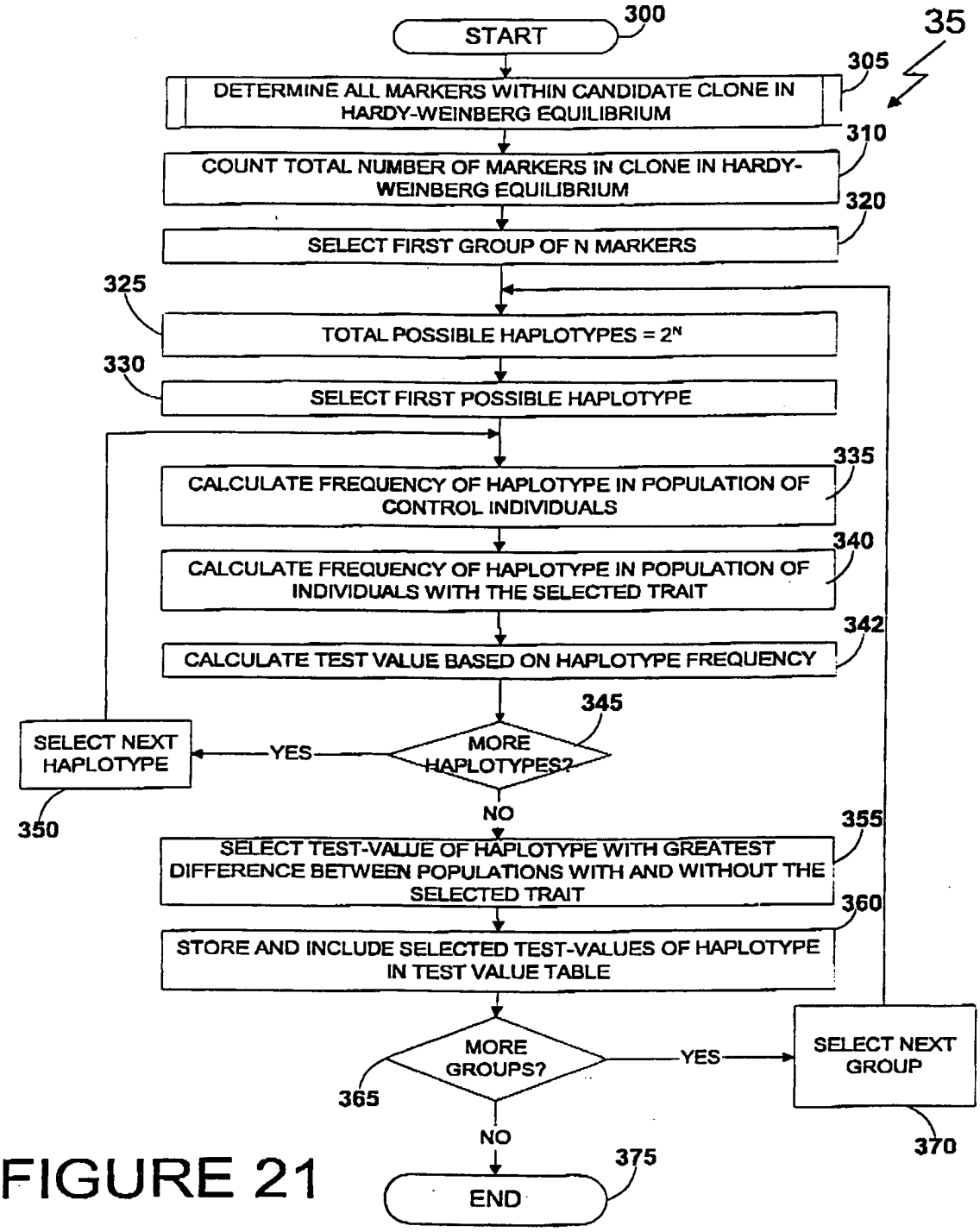



FIGURE 21

0838203 - 050504

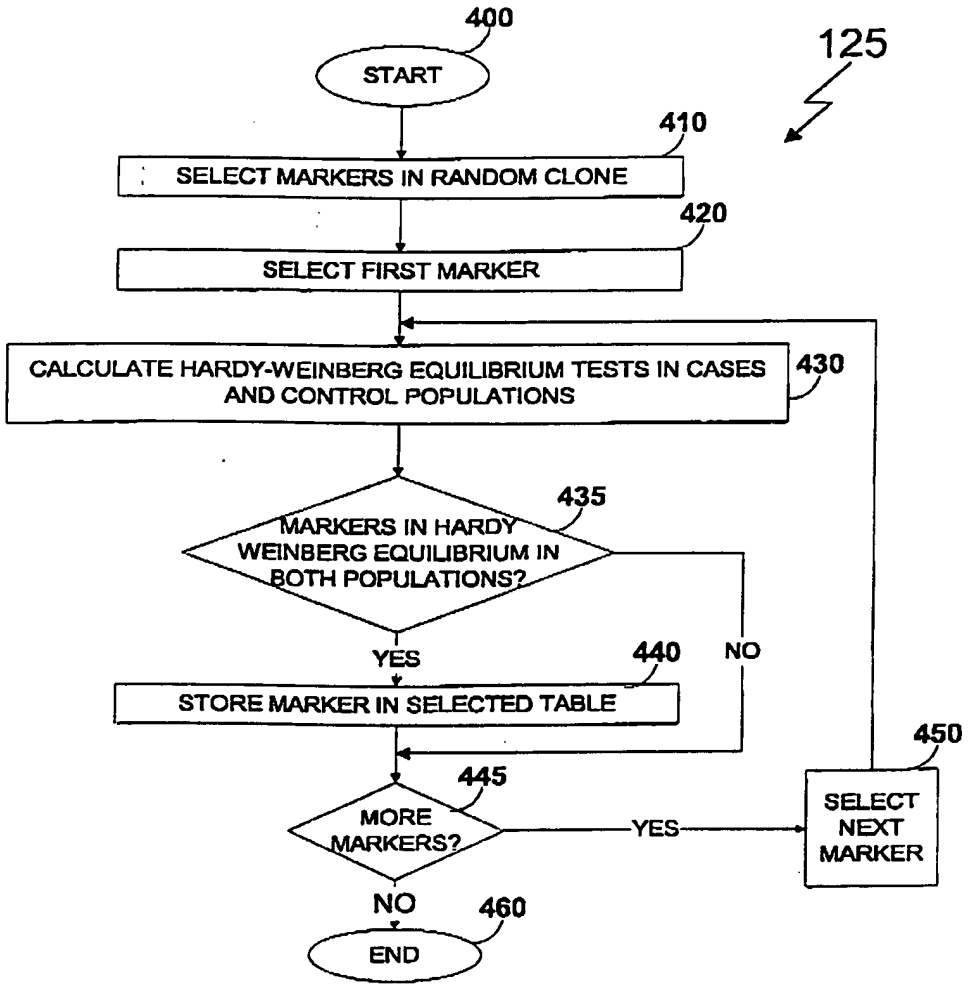



FIGURE 22

Docket No. 55.US4.DIV

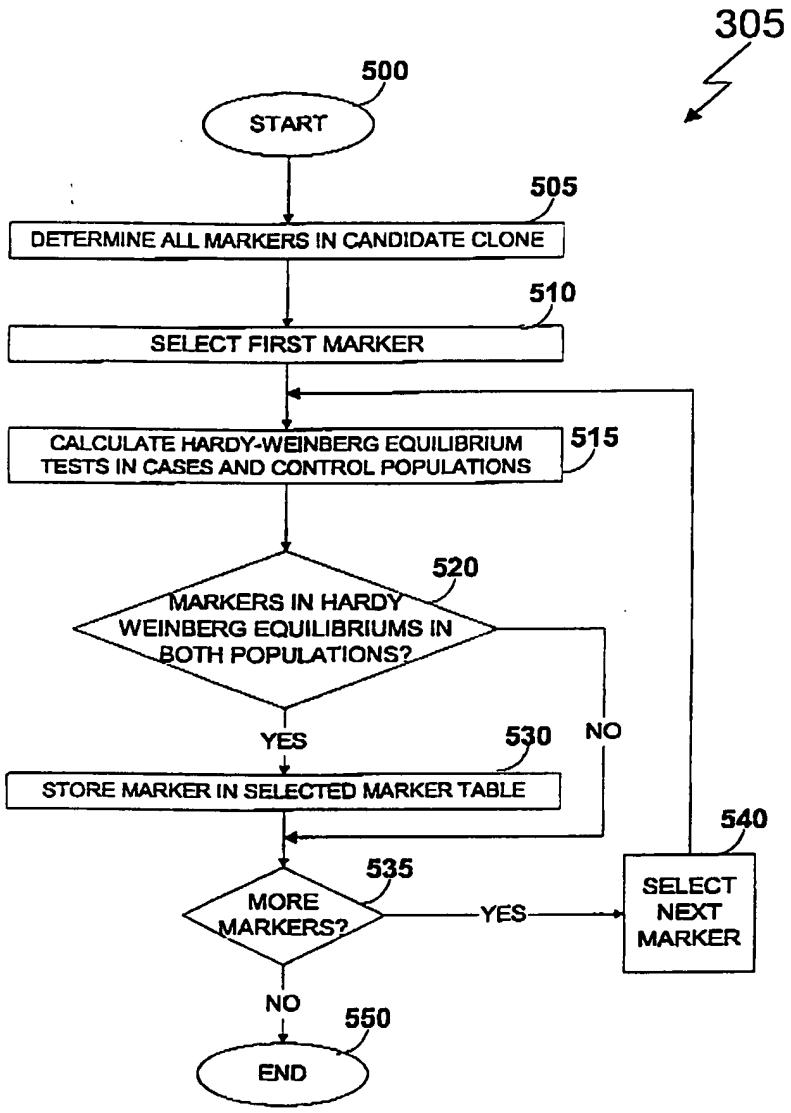
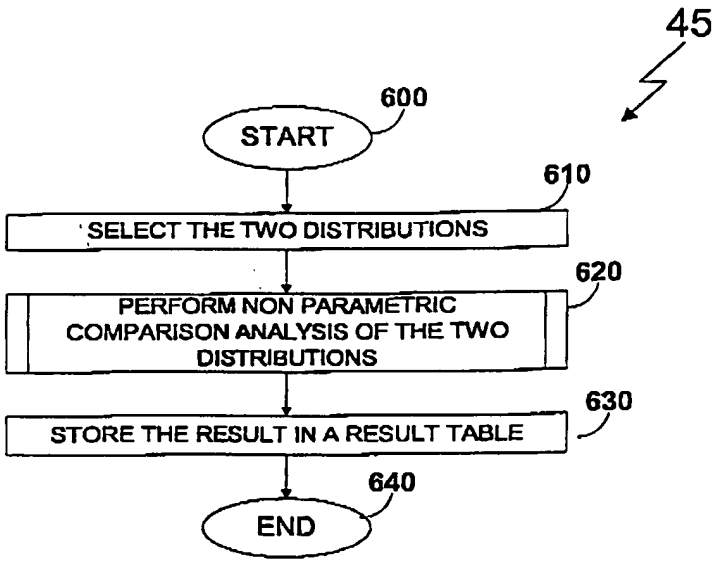



FIGURE 23

09505209 - 051501




09852809 - 051501

FIGURE 24

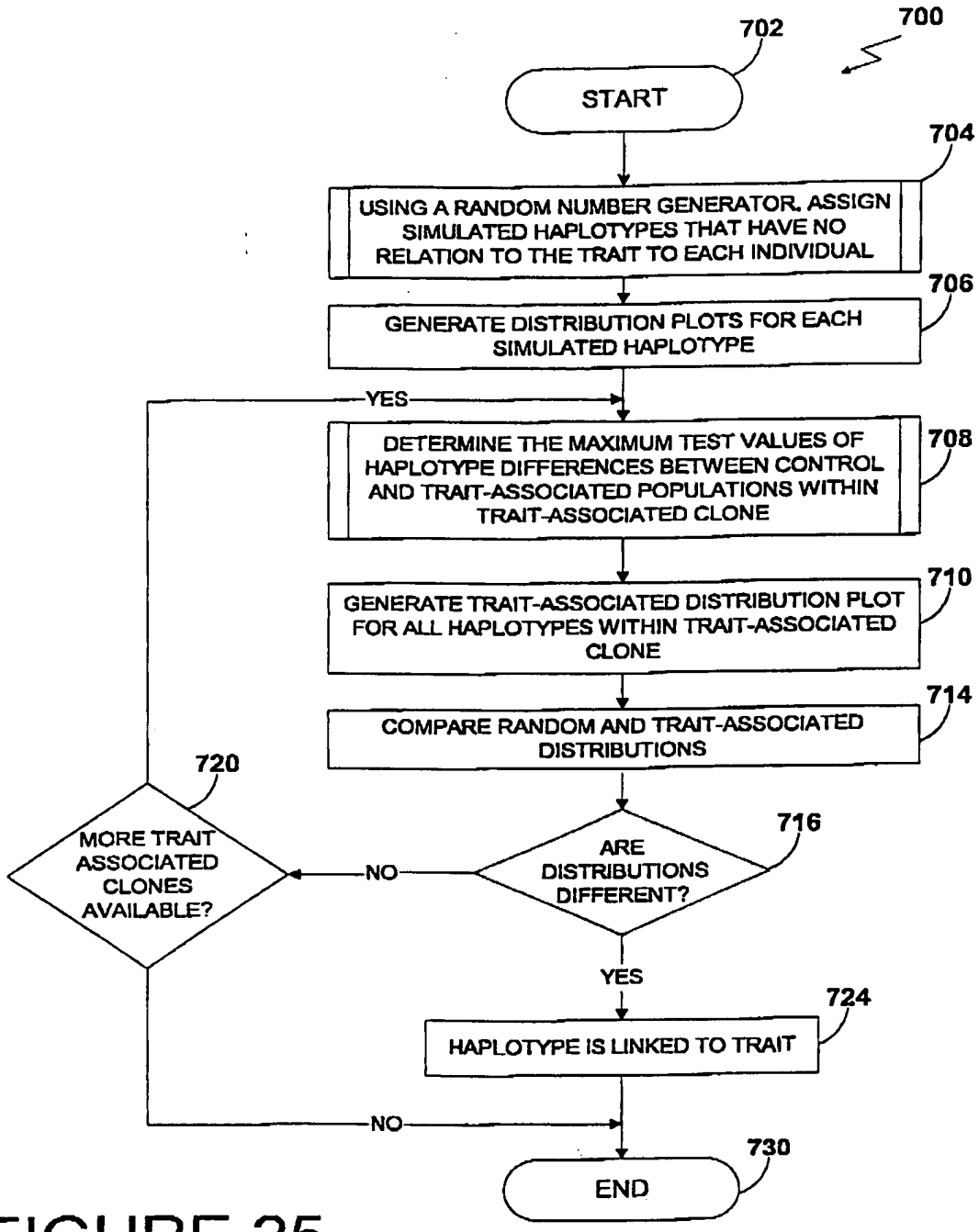



FIGURE 25

Docket No. 55.US4.DIV